## Description of algorithm to select "best" oligonucleotide adapter sequences.

Requirements for good sequences:

- Generates adequate hybridization signal intensity when employed in an experiment.
- Exhibits minimal cross-reactivity with other adapter sequences.
- Unique within the human genome sequence. This requirement can be extended to the genomic sequence of other organisms such as the fruit fly, the mouse, etc.

One method of generating sequences that meet the above requirements is to randomly generate sequences of given lengths and than pass these filters through a set of heuristic acceptance filters. In particular, the 24-mer Illumina Adapter sequences (IllumaCodes) were chosen as follows.

Generate a random 24-mer Computer randomly picks sequences from space oligonucleotide sequence of 24-mer oligonucleotide sequences. Abs $(T_m-T_{thres}) < 5$   $^{\circ}$ C Oligo Tm must be within a defined range. This Where  $T_{thres} = 72 \, ^{\circ}C$ requirment increases uniformity of hybridization signal.  $Abs(GG_{content} - GC_{thres}) < 0.15$ GC content must be balanced with AT content of oligo. This reduces sequence specific effects of hybridization. Self-Complementary Score < 14 Reduces oligonucleotide hairpin structures. A run of Gs or Cs can lead to poorly hybridizing  $GC_{run} < 5$ probes through G quartet structures. A high decoder-decoder complementarity score Decoder-Decoder Complementarity Score < 18 indicates that the two decoders are likely to cross-hybridize in solution. A high probe-decoder complementarity score Probe-Decoder Complementarity Score < 14 indicates that the decoder is likely to crosshybridize with the probe. The generated sequences are screened against a Blast against genome databases specified genomic database. (Unigene DB)

FIGURE 1

- -

## Flow Diagram for selection of probes sequences

Synthesize 768 oligo decoder sets on Oligator



Individually label decoders with two colors in HT format



QC: RP HPLC ~ 8 decoders/ plate



Pool decoders into 2 color -10 stage decode mix



Decode 768 bead series on array blocks in sets of 96 (by plate)



Select probe based upon decoding (good clustering) and normalization intensity

FIGURE 2

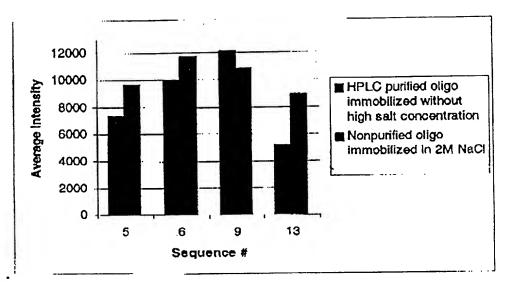
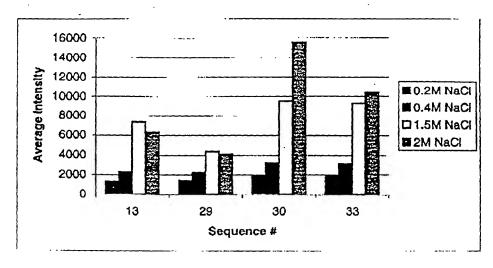


Figure 3



Figur 4

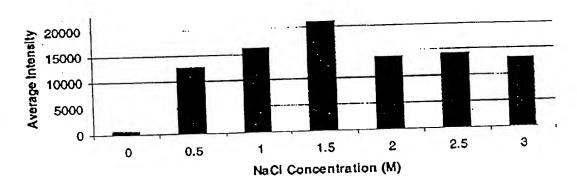


Figure 5